Maximum Likelihood Estimation

For EIE417 of SASTRA University (2024-25)

July 29, 2024

Introduction to Maximum Likelihood Estimation

- Maximum Likelihood Estimation (MLE) is a method to estimate the parameters of a statistical model.
- The principle behind MLE is to find the parameter values that maximize the likelihood function, which measures how likely it is to observe the given sample data given the parameters.
- If α is a parameter of a statistical model defined by the probability distribution function $P(X = x)$, then the likelihood of concurrence of happening of events $x_1, x_2, ..., x_n$ is

$$L(\alpha) = P(x_1 \cap x_2 \cap \ldots x_n|\alpha)$$

  If the events $x_1, x_2, ..., x_n$ are assumed to be independent then

$$L(\alpha) = P(x_1|\alpha)P(x_2|\alpha)\ldots P(x_n|\alpha) = \prod_{i=1}^{n} P(x_i|\alpha)$$

Introduction to Maximum Likelihood Estimation

- Likelihood Function

$$L(\alpha) = \prod_{i=1}^{n} P(x_i|\alpha)$$

- Likelihood Function

$$\ell(\alpha) = \log\left(L(\alpha)\right)$$

- Derivative

$$\frac{\partial \ell(\alpha)}{\partial \alpha} = \frac{\partial}{\partial \alpha}\left(\log\left(L(\alpha)\right)\right) = \frac{1}{L(\alpha)}\frac{\partial L(\alpha)}{\partial \alpha}$$

- Setting the derivative to zero

$$\frac{\partial \ell(\alpha)}{\partial \alpha} = 0 \Rightarrow \frac{\partial L(\alpha)}{\partial \alpha} = 0 \because L\left(\alpha\right) \neq \infty$$

Gaussian Distribution

- The probability density function is:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

- Given a sample $x_1, x_2, \ldots, x_n$, the likelihood function is:

$$L(\mu, \sigma^2 | x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | \mu, \sigma^2)$$

- The log-likelihood function is:

$$\ell(\mu, \sigma^2 | x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \log P(x_i | \mu, \sigma^2)$$

- Substituting the PDF:

$$\ell(\mu, \sigma^2 | x_1, x_2, \ldots, x_n) = -\frac{n}{2} \log(2\pi\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2$$

MLE for Gaussian Distribution

- Taking the partial derivative with respect to μ and setting it to zero:

$$\frac{\partial \ell}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (x_i - \mu) = 0$$

$$\Rightarrow \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

- Taking the partial derivative with respect to σ^2 and setting it to zero:

$$\frac{\partial \ell}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (x_i - \mu)^2 = 0$$

$$\Rightarrow \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{\mu})^2$$

Poisson Distribution

- The probability mass function is:

$$P(X = k|\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$$

- Given a sample $x_1, x_2, \ldots, x_n$, the likelihood function is:

$$L(\lambda|x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

- The log-likelihood function is:

$$\ell(\lambda|x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} (x_i \log \lambda - \lambda - \log x_i!)$$

MLE for Poisson Distribution

- Taking the partial derivative with respect to λ and setting it to zero:
$$\frac{\partial \ell}{\partial \lambda} = \frac{\sum_{i=1}^{n} x_i}{\lambda} - n = 0$$
$$\Rightarrow \hat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Binomial Distribution

- The probability mass function is:

$$P(X = k|n, p) = {}^nC_k p^k (1 - p)^{n-k}$$

- Given a sample $x_1, x_2, \ldots, x_n$, the likelihood function is:

$$L(p|x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} {}^nC_{x_i} p^{x_i} (1 - p)^{n-x_i}$$

- The log-likelihood function is:

$$\ell(p|x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} \left(\log\left({}^nC_{x_i}\right) + x_i \log p + (n - x_i) \log(1 - p)\right)$$

MLE for Binomial Distribution

- Taking the partial derivative with respect to p and setting it to zero:
$$\frac{\partial \ell}{\partial p} = \sum_{i=1}^{n} \left(\frac{x_i}{p} - \frac{n - x_i}{1 - p} \right) = 0$$
$$\sum_{i=1}^{n} x_i(1 - p) = \sum_{i=1}^{n} (n - x_i)p$$
$$\Rightarrow \hat{p} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Summary of MLE

- **Gaussian:**

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{\mu})^2$$

- **Poisson:**

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- **Binomial:**

$$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} x_i$$